Teaching Statement
Marinka Zitnik

My primary goal as an instructor is to teach students a strong, technical content that will enable students to make strides on both theoretical foundations and computational methodology. As a mentor, I aim to foster a can-do attitude, encourage students to tackle hard challenges and strive to train a next-generation of researchers capable of transforming our field.

On the road towards this long-term goal, I have actively sought opportunities to teach and advise students. I am active in sharing the knowledge, popularizing the field by presenting overview talks at Stanford and longer talks at conferences and workshops, and building bridges between different scientific communities. All these experiences have been fascinating and rewarding. I see teaching and advising students as essential elements of my future career as a professor.

Teaching and lecturing

I have had many opportunities to teach and interact with students. I was a teaching assistant for many courses in a computer science department (ranging from 30 to over 200 students), a teaching assistant for computer science courses in a medical school (ranging from 20 to over 100 students), and I have given a series of overview talks and tutorials at conferences.

I have shaped my teaching philosophy around a few core concepts. First, I value opportunities for hands-on learning. For example, a course might teach an algorithm, and I would like students to implement the algorithm and use it to analyze a dataset. I believe that students learn best when they apply abstract concepts in practice and are encouraged to find unconventional solutions. Second, I strive to improve the learning experience continuously. Along these lines, I regularly solicit feedback and use it to adjust the course content. Finally, to foster a can-do attitude I like to design a series of assignments that allow students to build and expand on newly acquired knowledge incrementally. I find that such an approach raises students' confidence in their skills and teaches students how to tackle hard problems by breaking them into small, manageable sub-problems.

Teaching and lecturing experiences

At Stanford, I had the opportunity to give guest lectures in the advanced undergraduate course on Social and Information Networks (CS224W), a class with over 200 enrolled students. I also designed engaging course assignments on biological networks and mentored student teams who applied their knowledge to analyze biomedical network datasets as part of the final course project. I was also involved in revamping the curriculum of the course (now called Analysis of Networks: Mining and Learning with Graphs), which included designing new lectures and refreshing existing ones. This experience was especially rewarding as I learned how to create lectures that match students' prior knowledge, how to organize the course so that students can build on their newly acquired knowledge, and how to best distribute technical content throughout the quarter.

At Baylor College of Medicine, I had a very different opportunity to co-teach a graduate-level course on Introduction to Data Mining (hosted by Prof. Gad Shaulsky and Prof. Adam Kuspa), a class with over 60 enrolled students from Quantitative and Computational Biosciences, Genetics, Genomics, and other graduate programs. I was responsible for designing and giving lectures as well as preparing assignments and teaching materials (e.g., handouts, data mining exercises using a visual programming paradigm). I also co-taught a course on Data Mining at The University of Pavia (hosted by Prof. Riccardo Bellazzi) for graduate students in Bioengineering, where I combined machine learning theory with hands-on data science exercises. These experiences have helped me become an effective teacher capable of adapting my teaching to my audiences.

Beyond this substantial experience in giving lectures in academic courses, I have given invited lectures, longer talks and tutorials at universities and major conferences. For example, I have given seminars in the Stanford Clinical Excellence Research Center (CERC; invited by Prof. Arnold Milstein) to teach the CERC fellows about the promise of data science to reduce healthcare costs and improve clinical care in the US. I gave tutorials on deep learning for biomedicine, integrative machine learning, and data fusion at the International Conference on Intelligent Systems for Molecular Biology (ISMB), Basel Computational Biology Conference (BC2), and the IEEE Engineering in Medicine and Biology Society (EMBC). It was a precious experience to compile these overview talks, which required distilling and synthesizing complex ideas in ways that could convey methodological and theoretical foundations in an intuitive but still precise manner. This is evidenced by the feedback from tutorial participants who sent me emails writing: “The materials were very well organized, loved the lectures. The speakers presented the math and the stats extremely clear. I also liked the fact that presenters periodically reminded us of what we had covered so far, and where we are going next. That really helped ground the tutorial.” I strive to introduce new material in ways that build bridges between different backgrounds people have, and that allows everyone interested in biomedical data science learn something new, which is also evidenced by the tutorial feedback: “It was very nice first to see the theoretical background and then the biological applications.”

Mentoring and advising

My mentors and collaborators have had a tremendous impact on the development of my research. As a mentor, I thus encourage students' growth, foster the research process, and use the scientific method as a framework for studying the world.

I enjoy mentoring and advising students, and I have seized many opportunities to mentor students and actively sought to improve...
my mentoring. I have had the opportunity to mentor a number of talented students in a wide variety of settings. For example, at Stanford, I mentored three students on their senior thesis projects. I mentored Monica Agrawal for her senior project on large-scale statistical analysis of disease pathways in biological networks, which led to a publication in a top-tier bioinformatics conference (PSB) with Monica as a first co-author. Monica did not have prior research experience; however, I held regular research meetings and worked together on formulating new problems, which helped her to make quick progress and eventually develop her own research ideas. Monica is also my co-author on a paper on graph neural networks, which we published in *Bioinformatics* and presented at a leading computational biology conference (ISMB). Monica has since joined MIT EECS as a Ph.D. student. Following this very positive mentoring experience, I am currently mentoring Evan Sabri Eyuboglu who has also produced new methodological results on his senior thesis project that will also lead to a publication.

I also mentored eight student teams (24 students in total) in an advanced project-based course on Mining Massive Datasets (CS341). Each group worked on a project related to data mining and machine learning algorithms for analyzing very large datasets. I held weekly meetings with each team and found that those meetings were essential in shaping the direction and progress of projects, formulating testable hypotheses, continually interpreting the results, and in maintaining positive relationships. I also mentored 12 undergraduate and M.Sc. research assistants (e.g., Priyanka Nigam’s work resulted in a tool for construction and analysis of large multimodal networks, mentoring of Agrim Gupta, Sheila Ramaswamy, and Viswajith Venugopal resulted in a construction of a multimodal biomedical network with over 2.3 billion edges).

I was also involved in summer undergraduate research internship programs (CURIS) at Stanford Computer Science (CS) where I mentored three students (e.g., Sagar Maheshwari’s project led to the Stanford Biomedical Network Dataset Collection). This experience was especially rewarding as the program aims to increase student diversity and encourage CS undergraduate students to get involved in research early in their careers. In addition to mentoring computer science and engineering students, I have also had the opportunity to advise outstanding graduate students in bioinformatics (e.g., Marija Durdevic at the Medical University of Graz, Balaji Santhanam at Baylor College of Medicine, and Camilo Ruiz at Stanford Bioengineering). When advising on these research projects, my first goal is to ensure that students conduct experiments that will quickly validate their ideas and give information on which direction they should head.

Through research rotations in our group, I additionally worked with a visiting Fulbright scholar and six junior Ph.D. students from a diversity of backgrounds (i.e., Computer Science, Statistics, Computational and Mathematical Engineering, and Biomedical Engineering). These collaborations resulted in a publication in top-tier computer science venue (KDD) and another two publications currently in preparation. I found that diversity of student skill sets represents a big challenge for mentoring and that this challenge is bigger in interdisciplinary research projects, such as method development for scientific fields like biology or medicine, where a basic understanding of the scientific field is often needed. Cultural differences between fields can in part explain this challenge; scientists from different areas often talk different languages and have different mindsets. To address this challenge, I like to learn about a student’s previous research experience and educational background to devise a project in which the student can thrive and make quick progress.

### Teaching interests

My research spans several areas of computer science, and I would be excited and qualified to teach any introductory-level undergraduate computer science course. Given my research and academic background, I would also be well-suited to teach a number of courses related to data science and machine learning, including Artificial Intelligence, Machine Learning, Network Science, Data Mining, and Deep Learning. I would also be delighted to teach courses in bioinformatics and computational biology, including Statistical and Machine Learning Methods for Genomics, Biological Modeling and Simulation, and Integrative Genomics.

In addition to teaching existing courses in these areas and potentially offering advanced graduate seminars in areas such as biomedical data science, I would like to create and design new courses at both introductory and advanced undergraduate levels that complement the existing course system. I sketch below two examples of such courses.

**Learning with Complex Networks.** This advanced undergraduate-level course will introduce students to core concepts from complex networks and data science, with a target audience of undergraduate students who are familiar with basic probability, linear algebra, and programming. The course will revolve around two main topics: data science and complex networks. It will complement graph theory and general data science classes by focusing on networks as a powerful modeling paradigm that can extract useful knowledge, understand a system’s dynamics, and make exciting discoveries. This will be an exciting course, as it will bring together ongoing research in machine learning and network science, combined with practical data science applications.

**Deep Learning in Healthcare and Biomedicine.** Recent breakthroughs in high-throughput measurement technologies and the explosion of detailed healthcare data are poised to make radical changes in healthcare and biomedicine, transforming areas such as diagnosis, genomics, surgical robotics, and drug discovery. In parallel, progress in deep learning is revolutionizing fields such as image recognition, natural language processing and, more broadly, artificial intelligence. This course will explore the exciting intersection between these two advances. This graduate-level course will cover the ongoing developments in deep learning (supervised, unsupervised, and generative models) with the focus on applications of these methods to healthcare and biomedical data, which are beginning to produce dramatic results. In addition to predictive modeling, the course will emphasize how to visualize and extract interpretable, biomedical insights from deep models.